

# **EXHIBIT F**

505 Montgomery Street, Suite 2000  
San Francisco, California 94111-6538  
Tel: +1.415.391.0600 Fax: +1.415.395.8095  
www.lw.com

**LATHAM & WATKINS** LLP

FIRM / AFFILIATE OFFICES

Austin	Milan
Beijing	Munich
Boston	New York
Brussels	Orange County
Century City	Paris
Chicago	Riyadh
Dubai	San Diego
Düsseldorf	San Francisco
Frankfurt	Seoul
Hamburg	Silicon Valley
Hong Kong	Singapore
Houston	Tel Aviv
London	Tokyo
Los Angeles	Washington, D.C.
Madrid	

May 7, 2024

**VIA EMAIL**

Zach Savage  
zsavage@susmangodfrey.com  
Susman Godfrey L.L.P.  
One Manhattan West  
New York, NY 10001-8602

Re: *The New York Times Company v. Microsoft Corporation, et al.*, Case No.: 23-cv-11195-SHS

Counsel:

I write regarding The New York Times Company's ("NYT") responses and objections to OpenAI's First Set of Requests for Production. Please provide your availability to meet and confer about the items discussed below.

**A. Request Nos. 2, 7, 20, 21, and 23 (Regurgitation Prompting)**

NYT objects to a number of requests relating to its efforts to cause the GPT Services to regurgitate NYT articles, on grounds of work-product doctrine. NYT Resp. to Request Nos. 2 (attempts to reproduce NYT works using GPT Services), 7 (NYT's OpenAI accounts), 20 (creation of Exhibit J), 21 (OpenAI accounts used to create outputs in complaint), and 23 (process for obtaining outputs cited in complaint).

As a threshold matter, NYT has no claim of work-product protection over data (including, *e.g.*, model and system prompts) it or its agents have submitted to OpenAI via the GPT Services. These submissions to a third party are not confidential, particularly in light of GPT Services users' express assent to terms explicitly reserving OpenAI's right to use such data to defend itself in litigation. *See, e.g.*, OpenAI, Privacy Policy § 2. We are not aware of any authority suggesting that a party can claim work product protection over information submitted directly to its anticipated litigation adversary. To the extent NYT is relying on any such authorities, please provide them. In any case, NYT waived work-product protection over any materials relating to its efforts to cause the GPT Services to allegedly regurgitate data by explicitly relying on selected results of those efforts in its complaint. *See In re Commodity Exch., Inc., Gold Futures & Options Trading Litig.*, No. 14-cv-2548, 2019 WL 13046984, at \*2–4 (S.D.N.Y. Feb. 25, 2019). To the extent NYT has authorities that suggest otherwise, please provide them.

**LATHAM & WATKINS** LLP

Each of these requests is also highly relevant to NYT's claims, which hinge in substantial part on NYT's allegation that "individuals can access [NYT] content through Defendants' [] products without having to pay for it." Compl. ¶ 157. OpenAI, for example, is entitled to documents regarding the scope and extent of NYT's efforts to cause the GPT Services to allegedly regurgitate NYT articles—along with any failed attempts to do so—because those documents may demonstrate that users can neither "access" NYT content through those services as a general matter nor without substantial effort. Similarly, OpenAI is entitled to information sufficient to identify any OpenAI accounts NYT or its agents used during those efforts, which may enable OpenAI's independent discovery of similarly relevant information in its own possession. *See* Request Nos. 7 & 21.

Relatedly, Request No. 2 seeks "[a]ll Documents and Communications regarding any attempt by You, including failed attempts, to reproduce any of Your Published Works via GPT services." In its response, NYT states that it "reasonably construes this Request to refer to [its] process for obtaining the GPT Services outputs cited in the Complaint." NYT Response to Request No. 2. But the scope of Request No. 2 is broader: it encompasses documents arising out of any attempts by NYT to reproduce the asserted works via GPT Services—whether or not the resulting outputs are referenced in the complaint—including user logs, screenshots, spreadsheets, or any other records of outputs created in the course of those attempts, regardless of whether those outputs actually contain expressive material from the asserted works.

As we informed you by our letter of February 9, 2024, NYT had an obligation to preserve all documents relating to its efforts to cause the GPT Services to allegedly regurgitate training data. Please confirm that NYT will search for and produce all non-privileged documents within the scope of these requests. To the extent NYT plans to withhold documents on grounds of attorney-client privilege or work-product doctrine, please confirm that NYT will provide a privilege log that identifies any withheld materials.

### **B. Objections Based on Reporters' Privilege**

NYT objects to other requests on the basis of "reporters' privilege pursuant to the First Amendment of the U.S. Constitution or the New York Shield Law, N.Y. Civ. Rights § 79-h." NYT states that it will "not search for or produce Documents or Communications protected by the reporters' privilege." NYT Responses to OpenAI Request Nos. 1, 3, 7, 10, 11, 12, 16, 17, 22, 24, 47, 48, 49, 50, 56. To the extent NYT is relying on authorities suggesting that the reporters' privilege excuses the owner of a news publication from searching for documents in response to discovery requests in a litigation brought by the news publisher and claiming harm to its reporting activities, please provide those authorities and a description of the categories of documents you plan to exclude from your search. Failing that, please confirm you will not exclude materials from the scope of your search based on reporters' privilege and will provide a privilege log of all documents you intend to withhold on the basis of that alleged privilege.

### **C. Objections Based on Common Interest Privilege**

NYT objects to various requests "to the extent that [they] seek[] material protected by the ... common interest [doctrine]." NYT Response to OpenAI Request Nos. 1, 3, 5, 7, 8, 10, 11, 13—

26, 33, 35–38, 40, 41, 43, 45, 52, 54–57. Please identify the third party or parties with whom NYT asserts a common interest and describe the basis for NYT’s assertion of the doctrine. To the extent NYT intends to withhold documents on the basis of the common interest doctrine, please confirm that NYT will identify those materials on a privilege log.

#### **D. Request Nos. 4–6 (Other Generative AI Services)**

NYT contests the relevance of requests about the outputs of “Generative AI services other than GPT Services” and refuses to produce responsive documents. NYT Resp. to OpenAI Request Nos. 4 (reproduction, display, or distribution of NYT works), 5 (NYT’s attempts to cause regurgitation), 6 (outputs that summarize or reference NYT works). This evidence is relevant to several claims and allegations in the complaint, including NYT’s assertion that “GenAI” services “threaten high-quality journalism,” and its allegation that “given the right prompt, [generative AI services] will repeat large portions of materials they were trained on.” Compl. at 14 & ¶ 80. It is also relevant to a number of other issues, including the behavior of OpenAI’s services as compared to other LLMs (relevant to fair use) and the extent to which users account for the known phenomenon of hallucination when interpreting AI outputs. *See* Compl. ¶ 142. As another example, NYT’s awareness of whether other generative AI companies have engaged in the same conduct NYT contends to be copyright infringement here and NYT’s inaction with respect to those companies would undermine NYT’s claim for injunctive relief, as it would indicate a lack of irreparable harm from the conduct NYT alleges.

At a minimum, documents sufficient to identify any other generative AI services you believe were trained on the asserted works are relevant to NYT’s claim that OpenAI’s conduct directly caused NYT to suffer financial harm, along with the extent of harm attributable to that conduct. Please confirm that NYT will produce non-privileged documents responsive to these requests and provide a privilege log that identifies any materials withheld on grounds of attorney-client privilege or work-product doctrine.

#### **E. Request Nos. 8–12 (Creation of Asserted Works)**

Requests No. 8 through 12 seek information relating to the preparation and scope of copyright protection of the works NYT asserts in this lawsuit. *See* Request Nos. 8 (allegations of infringement), 9 (allegations of plagiarism), 10 (expressive, original, and human-authored content), 11 (non-expressive, non-original, or non-human-authored content), 12 (reference material). NYT has refused to produce documents in response to these requests on grounds of reporters’ privilege, *see supra* (addressing this alleged privilege), and further contests the relevance of Requests 8, 9, and 12. The only documents NYT has agreed to produce in response to these requests are the asserted works themselves. *See* NYT Resp. to OpenAI Requests 10 and 11.

Requests 8, 9, and 12 are relevant to NYT’s claims and OpenAI’s defenses in several respects. For example, NYT cannot assert copyright infringement claims based on OpenAI’s alleged reproduction of portions of NYT articles that were “copied from other [third-party] works,” like reference material or pre-existing articles. *Feist Publ’ns, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 345 (1991) (works not protectable if “copied from other works”); *see also Estate of Burne*

*Hogarth v. Edgar Rice Burroughs, Inc.*, 342 F.3d 149, 166–67 (2d Cir. 2003) (copyright registration creates rebuttable presumption of validity that defendant may challenge with “other evidence”). So too for material without a human author. *Thaler v. Perlmutter*, No. 22-cv-1564, 2023 WL 5333236, at \*7 (D.D.C. Aug. 18, 2023). And evidence that NYT created the asserted works by summarizing works of third-party journalism is directly relevant to the validity of NYT’s assertion that the GPT Services allegedly violate New York misappropriation law by doing exactly the same thing. *See, e.g.*, Compl. ¶ 194.

Please confirm NYT will search for and produce non-privileged documents responsive to these requests and log documents withheld on privilege grounds. At a minimum, these requests require NYT to produce, for each of the asserted works, (1) any documents regarding allegations that the work infringes or plagiarized other works; and (2) any materials that “informed the preparation” of each of the asserted works, including without limitation reference materials gathered during the course of each work’s creation in NYT’s custody and possession.

#### **F. Request Nos. 13–15 (Ownership and Registration of Asserted Works)**

Requests No. 13 through 15 seek information relating to NYT’s ownership and registration of the asserted works. *See* Request Nos. 13 (ownership disputes), 14 (correspondence with Copyright Office), 15 (authorship and work-for-hire agreements).

*Request No. 13.* NYT has refused to produce any documents in response to Request No. 13. Documents regarding ownership disputes are relevant to the extent they suggest that the NYT is not the legal or beneficial owner of the works it asserts in this lawsuit. 17 U.S.C. § 501(b). OpenAI is willing to limit Request No. 13 to exclude DMCA Takedown Notices relating to user-submitted content. *See* NYT Resp. to OpenAI Req. No. 13 (objecting to relevance of DMCA takedown notices that “apply to user-submitted content”). Subject to that narrowing, please confirm that NYT will produce responsive, non-privileged documents in response to this request (and log documents withheld on privilege grounds) or explain NYT’s apparent position that NYT’s ownership of its asserted works is not relevant to its claims.

*Request No. 14.* NYT has refused to produce its correspondence with the Copyright Office (other than its deposit copies) in response to Request No. 14. NYT’s correspondence with the Copyright Office as to each of the asserted works is relevant to the extent such documents reveal any questions raised by the Office as to the registrability of those works, and any responses to those questions by NYT. Merely producing NYT’s deposit copies is not sufficient to satisfy the scope of this request. Please confirm NYT will produce responsive documents within the scope of this request.

*Request No. 15.* NYT has refused to produce documents relating to its ownership of the asserted works other than “agreements . . . related to authorship” in response to Request No. 15. That is unacceptable. At a minimum, this request requires NYT to provide, for each of the asserted work, (1) if the work is a work made for hire, the relevant employment or contract pertaining to the creation of the work by the person(s) who created it under 17 U.S.C. § 201(b), or (2) if the work is not a work made for hire, signed agreements pertaining to the transfer of rights in that work from its original author to NYT under 17 U.S.C. § 204(a), including any intermediate

transfers. These documents are relevant to the threshold question of whether NYT has standing to bring infringement claims based on those works. *See, e.g., John Wiley & Sons, Inc. v. DRK Photo*, 882 F.3d 394, 410–15 (2d Cir. 2018). Please confirm NYT will produce non-privileged responsive documents and provide a log describing any documents withheld on privilege grounds.

**G. Request Nos. 24 (Correspondence with Other Plaintiffs)**

Request No. 24 seeks communications between NYT and other news, media, or writers' organizations and publishers, as well as other named plaintiffs in certain lawsuits against OpenAI, related to copyright and artificial intelligence. NYT refused to produce documents responsive to this request and argues that the documents it seeks are not relevant.

NYT's statements about copyright and artificial intelligence to third parties are relevant to a variety of issues the NYT has affirmatively raised in this case. NYT alleged, for example, that generative AI "threaten[s] high-quality journalism," imperils the "traditional business models" of NYT "and its peers," and create an "enormous" "cost to society." Compl. at 14. It has also alleged that those services "deprive [NYT] of subscription, licensing, advertising, and affiliate revenue." Compl. ¶ 5. OpenAI is entitled to evaluate the validity of those statements, including by investigating whether NYT has said otherwise in non-privileged correspondence with other media companies and similar entities. OpenAI is also entitled to investigate the extent to which NYT has engaged in anticompetitive horizontal conduct with its direct competitors, which would support, *inter alia*, OpenAI's anticipated defenses of unclean hands and copyright misuse. *See Primetime 24 Joint Venture v. Nat'l Broad. Co., Inc.*, 219 F.3d 92, 102 (2d Cir. 2000).

Please confirm that NYT will produce non-privileged documents responsive to this request and, to the extent NYT withholds documents on grounds of common interest (or any other privilege or doctrine), that NYT will log all such documents.

**H. Request Nos. 33 and 43 (Licenses and Attempts to License)**

Request No. 33 seeks all documents and communications relating to licenses or attempts to license the asserted works, and Request No. 43 seeks documents and communications relating to commercial arrangements by which third parties reproduce, distribute, display, or perform the asserted works. NYT objects to their relevance and proposed to narrow this request to "documents sufficient to show the key terms of relevant licenses active from January 1, 2018 to the present." *See* NYT Resp. to Req. Nos. 33 and 43. This would appear to exclude broad categories of responsive materials, including even potentially the license agreements themselves. Please explain NYT's basis for refusing to produce such materials.

**I. Request Nos. 50 and 51 (NYT's AI Tools)**

Request No. 50 seeks "All Documents and Communications relating to the training of Your artificial intelligence satellite images tool referenced in the February 19, 2024 Reuters Institute article, 'New York Times publisher A. G. Sulzberger: "Our industry needs to think bigger"' including without limitation any base models and training data You used." Request No. 51 seeks "All Documents and Communications relating to the training of Your artificial intelligence tool



referenced in the February 20, 2024 Axios article, ‘Exclusive: NYT plans to debut new generative AI ad tool later this year’ including without limitation any base models and training materials You used.” Based on objections to relevance and privilege, NYT refuses to produce any documents responsive to these requests.

Documents relating to NYT’s own training and use of artificial intelligence models are highly relevant to issues NYT has affirmatively raised in this litigation, including, for example, NYT’s assertion that generative AI “threaten[s] high-quality journalism.” Compl. ¶ 47. NYT’s own use of artificial intelligence models, and its use of those models to advance its journalistic efforts, may yield admissions regarding the transformative nature of generative AI models. Please confirm that NYT will produce documents responsive to this request.

#### **J. Request Nos. 54 and 55**

This request seeks “All Documents and Communications relating to The New York Times English Gigaword (LDC2003T05) of the Linguistic Data Consortium of the University of Pennsylvania, including but not limited to all articles included therein, any licensing arrangements, and any negotiations between You and the Linguistic Data Consortium or University of Pennsylvania.” Similarly, Request No. 55 seeks “All Documents and Communications relating to The New York Times Annotated Corpus (LDC2008T19) of the Linguistic Data Consortium of the University of Pennsylvania, including but not limited to all articles included therein, any licensing arrangements, and any negotiations between You and the Linguistic Data Consortium or University of Pennsylvania.” NYT objects on relevance and privilege grounds and offers to narrow this request to “documents sufficient to show the content of the corpus.” NYT Resp. to OpenAI Req. Nos. 54 and 55. NYT, in other words, refuses to produce licensing arrangements relating to these corpora.

Those documents are highly relevant to OpenAI’s anticipated fair use defense and, in particular, the extent to which there exists a “potential [licensing] market” for use of the asserted works to train generative AI models. 17 U.S.C. § 107(4). The scope of relevant information includes, for example, the terms of every licensing arrangement between the NYT and a third-party regarding these corpora and the remuneration NYT has earned from any negotiated agreements (if any). Please confirm that NYT will search for and produce non-privileged documents responsive to this request.

#### **K. Requests for Clarifying or Narrowing Scope**

In response to various requests for production (e.g., Request Nos. 18, 30, 35, 37, 42, 45), NYT requests that the scope of the request be clarified or narrowed. We are happy to meet and confer about the scope of these requests. Please be prepared to discuss which aspects of these requests NYT contends are unclear or overbroad.

LATHAM & WATKINS<sup>LLP</sup>**L. Search Protocol**

In response to various requests, NYT indicates that it will agree to produce documents located after a search “conducted pursuant to an agreed-upon search protocol.” Please provide a proposal for discussion.

**M. Objections to Definition of “Employee”**

NYT objects to the definition of “Employee” and appears to take the position that it should be interpreted to exclude officers, directors, and former employees. Relevant materials in NYT’s possession are not shielded from discovery merely because they were created or received by a former employee or a current officer or director, and the basis for NYT’s objection is not clear. If there are any particular categories of documents that NYT intends to exclude from the scope of its search or production based on this objection, please identify them and explain the justification. Otherwise, please confirm that NYT will not withhold any documents otherwise responsive to OpenAI’s requests on the basis of this objection.

Sincerely,

/s/ Joe Wetzel

Joe Wetzel  
of LATHAM & WATKINS LLP